

Full software support for game learning analytics*

Antonio Calvo-Morata^{1,*†}, Julio Santilario-Berthilieres^{1,†}, Cristina Alonso-Fernández^{1,†},
Manuel Freire^{1,†}, Iván Martínez-Ortiz^{1,†} and Baltasar Fernández-Manjón^{1,†}

¹ Software Engineering and Artificial Intelligence Department, Complutense University of Madrid, Spain

Abstract

Learning Analytics (LA) continue to be successfully applied in multiple domains. For serious games in particular, Game Learning Analytics (GLA) aim to gain insights into player actions and behaviors from player interaction data. Such insights can provide the different stakeholders (e.g., serious game designers and developers, teachers, educators) actionable information about the learning process of players as well as evidence-based information about the learning design of games. Despite its potential, the implementation of GLA continues to be complex, costly, technology-dependent, and in many cases a fragile process that is difficult to conduct for game developers. Most of the current GLA approaches and platforms are ad-hoc proprietary solutions which are difficult to reuse or adapt to different technologies or contexts. To address these issues, we propose a standard-based open-source platform with a set of tools that simplify not only the implementation of GLA in serious games but also their experimental design and validation. Our open-source platform, called SIMVA (SIMple VAlidator), uses the international IEEE xAPI standard to collect player interaction data. SIMVA provides software for streamlining GLA, including player tracking (with trackers currently available for Unity and JavaScript games), player interaction data storage, and basic data analysis and visualization. The use of the xAPI standard increases the reusability of GLA data and allows stakeholders to create, reuse, or modify the different components of the system while keeping full ownership of GLA data. This includes different GLA data storage systems (called Learning Record Stores) and trackers that can be reused, or implemented anew for different game platforms. SIMVA also allows the configuration of different visualizations, making it easy to create graphics that display data collected from games. SIMVA is also a GLA research platform for game validation and impact measurement. It supports different experimental designs (e.g., with a control group, A/B testing) and allows data collection from pre-post questionnaires, which are the still most common instrument to measure the impact of serious games on players. These questionnaires are also linked to player interaction data via an anonymized code, which is also supplied to games. Use of anonymous linked identifiers allows SIMVA to combine in-game interaction data with qualitative information from questionnaires to help assess the impact of the game on players, in a way that complies with European data privacy regulations, and specifically the EU GDPR. This approach offers the opportunity to turn serious games into scientifically validated learning tools, although it requires a strong experimental design and advanced software infrastructure. By using SIMVA, researchers and game developers can enhance the effectiveness of games by improving learning outcomes and generating credible evidence of their educational value. This in turn can impact the application of games in educational scenarios, as their formal validation will increase the confidence of educators and institutions in these educational tools.

Keywords

learning analytics, serious games, technology-enhanced learning

1. Introduction

The information gathered from Learning Analytics (LA) [1] in different learning scenarios has successfully provided stakeholders with information to improve their teaching and learning experiences. This is also the case for serious games, where Game Learning Analytics (GLA) [2] uses player interaction data to gain insights into player actions and behaviors. With serious games having

*LASI Spain 25: Learning Analytics Summer Institute Spain 2025, Vitoria-Gasteiz, May 26-27, 2025

^{1*} Corresponding author.

[†] These authors contributed equally.

✉ acmorata@ucm.es; jsantila@ucm.es; calonsofernandez@ucm.es; manuel.freire@fdi.ucm.es; imartinez@fdi.ucm.es; balta@fdi.ucm.es

>ID 0000-0001-8701-7582; 0009-0003-2520-507X; 0000-0003-2965-3104; 0000-0003-4596-3823; 0000-0001-6595-5690; 0000-0002-8200-6216

Copyright © 2025 for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

been applied in multiple domains with proven effectiveness, the incorporation of analytics provides more evidence-based information to support and guide their design, development, application and deployment in real learning environments. The insights obtained from GLA can provide different stakeholders (e.g., serious game designers and developers, teachers) actionable information about the learning process of players as well as evidence-based information about the learning design of games [3].

While new serious games are constantly being developed, their wider adoption still faces several issues. First, in most learning environments, teachers lack control while their students are playing a serious game, and need actual and near real-time information to be able to keep track of their students' actions to direct the class [4]. There is even a more relevant issue: many serious games lack validation, so there is no proof that they help to meet the goal they are intended to, such as increasing learning, awareness, or changing attitudes [5]. Game validation, when carried out, is still mostly performed through traditional outside measurements, such as questionnaires, which fail to capture actual interactions with the game [6]. For that purpose, GLA can be a powerful tool to gather evidence to better understand the actions of students within serious games and improve their learning process [7].

In addition to the evidence-based validation of serious games, the use of GLA within serious games can be used for multiple purposes to benefit all stakeholders, of which teachers are only a salient example. Educational institutions could be more willing to integrate serious games into their activities when provided actual evidence of their effectiveness [8]. Students, who actually play the games to learn, would also find information about their progress, areas of improvement, or errors made very relevant, both while playing and after finishing the activity. Finally, serious game designers, developers and researchers could improve the learning and game design of serious games to adapt it to players requirements and obtain data about games effectiveness and applicability in classes [9].

Despite this potential, the implementation of GLA continues to be a complex, costly, technology-dependent, and in many cases fragile process that is difficult to conduct for many game developers. Traditional Game Analytics is usually carried out by enterprises on closed-source, proprietary systems such Wooga, Zynga, Microsoft, or Ubisoft, which do not address the learning domain, making their offerings difficult to reuse or adapt to different technologies or contexts. On the other hand, most of the current LA approaches and platforms are ad-hoc proprietary solutions, which also results in their being difficult to reuse or adapt to different technologies or contexts. Some open-source platforms and tools to collect and analyze LA data from learning scenarios do exist, including EXCALIBUR-LA [10], an extendable and scalable infrastructure build for Learning Analytics with Moodle logs; Tutor Shop [11], a component used to sequence tutor problems in a web-based environment with LA support; moodle2eventLog [12], a tool to transform raw Moodle logs into enriched, analysis-ready event logs; or xAPI Made Easy [13], a Learning Analytics infrastructure for interdisciplinary projects.

The lack of free and open-source GLA solutions also limits their application to educational scenarios, which often lack the possibility to obtain software licenses. Even those available most often lack supporting tools for teachers to understand how to integrate a learning activity using a serious game into their classes. This lack of tools also results in a barrier to the application of serious games and their evaluation [14]. To address these issues, we propose a full software support system for game learning analytics: an open-source platform composed of several interchangeable modules that, together, cover the requirements of all stakeholders involved in the process.

2. Proposal: GLA support system

The GLA support system proposed, SIMVA (SIMple VAlidator) [15], aims to cover most of the identified issues when applying serious games in actual contexts. For that, we have designed and developed a full system that covers the management of users and learning activities, the collection of information from serious games and its storage, and the analysis and visualization of that

information. All the GLA system modules that compose SIMVA are open-source and available for free on GitHub.

To develop such a complex and varied system, the use of a standard data format to gather players' actions in games was essential. For that purpose, we use the IEEE Experience API (xAPI) standard and its profile for Serious Games, xAPI-SG [16]. This standard allows us to collect player interaction data by defining a default set of verbs (actions), activity types and extensions that are common in learning activities and in serious games. The standard uses a JSON-based format with 3 main fields: an actor (who performed the action), a verb (the action performed) and an object (which received the action). For instance, Figure 1 presents an example xAPI-SG data trace (also called a *statement*) stating that a user (identified only by its anonymous actor identifier, "fybu") completed a game level. Additional information can be collected using other fields of the trace, for instance: the game level completed, the result of the action (e.g., score achieved in the level, progress reached in the game, etc.), the context of the action, and the time when the action occurred (collected in the timestamp field).

```
{
  "actor": {
    "account": {
      "homePage" : "...",
      "name" : "fybu"
    }
  },
  "verb": {
    "id": "http://adlnet.gov/exapi/verbs/completed"
  },
  "object" : {
    "definition" : {
      "type" : "https://w3id.org/xapi/seriousgames/activity-types/level"
    },
    "id" : "..."
  },
  "result" : ....,
  "context": ....,
  "id": "ad48cb27-0536-4b22-a0fb-6c4f80c57d5a",
  "timestamp" : "2021-06-09T19:32:57.369Z",
  "stored" : "2021-06-09T19:41:23.456Z"
}
```

Figure 1: Example of an xAPI-SG statement, highlighting the main sections.

The use of the xAPI standard to represent player interactions in games allows all modules of our full support system to have a common shared language to communicate with each other. The different modules can be created from scratch, reused, or modified; for example, by creating new trackers for different game platforms.

SIMVA, as a GLA support system, is composed of 4 main modules (Figure 2): (1) users and activities management, (2) telemetry, (3) persistence, and (4) analysis and visualization.

Through these modules, Simva provides different features that can be of interest to the different stakeholders:

- For **teachers**, it supports the application of games in their classes by collecting and analyzing student data that can provide them with insight into progress and results of the activity. SIMVA allows teachers to give access to specific users (students) to each study/activity, and to view which users have completed each activity – while restricting access to the data to those who have permission.
- For **students**, it enhances gameplay by collecting information about their results that can be used both to provide them with feedback and to improve game design.
- For **researchers**, it provides tools for validating educational games and measuring their impact, by allowing the creation of experimental designs where data are collected from different sources, such as questionnaires and game interaction data.

- For **game designers and developers**, it provides evidence to analyze and improve their games.
- For **educational data scientists**, it provides a framework for exploring in-game behavior data.

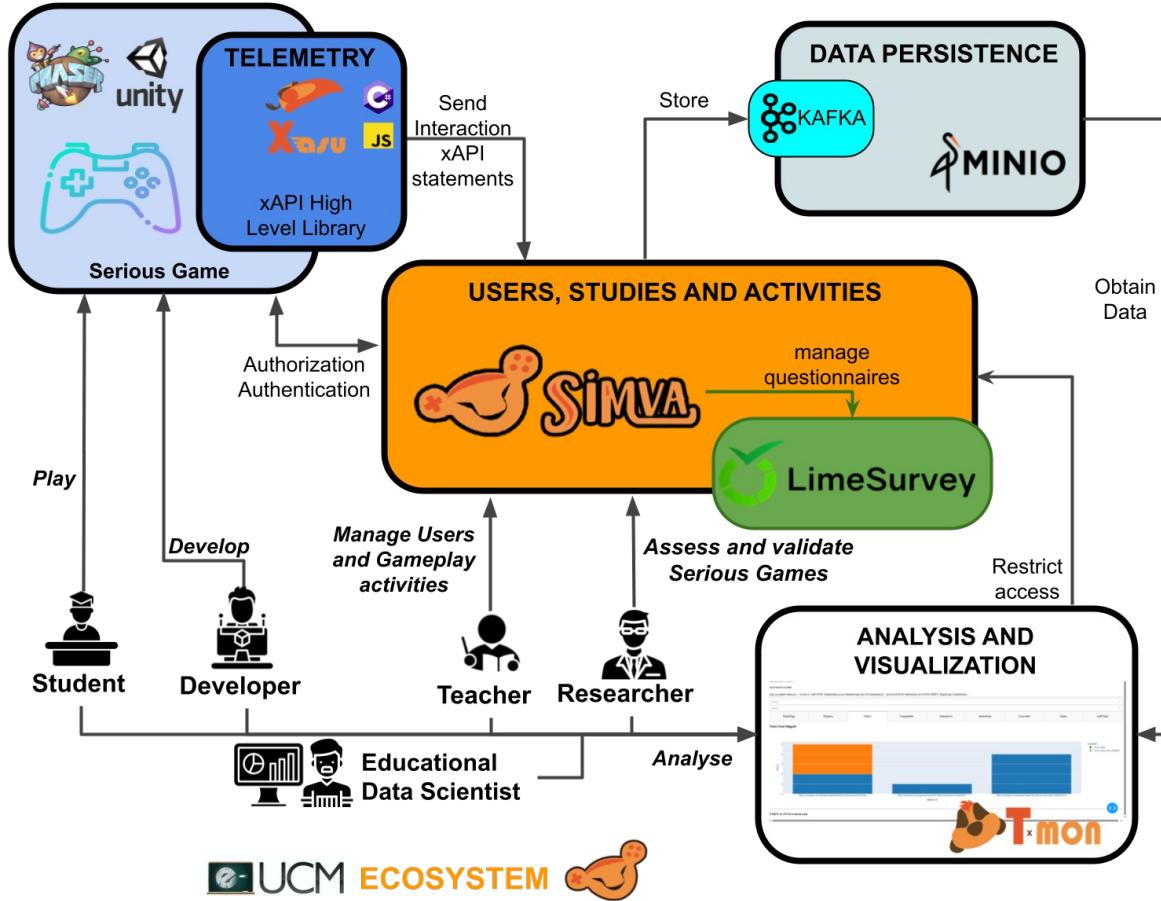


Figure 2: Modules of the SIMVA GLA support system (center) within the e-UCM ecosystem.

2.1. Users and activities module

This module is the core of the system and allows users to manage and configure different learning activities. This includes the creation of studies (aggregations of activities) and activities, the configuration of groups to participate in the activities of a study, and the different setups for experiments. Activities can include game-based learning activities (i.e., in connection with the telemetry module) and/or questionnaires. Figure 3 shows an example of a study consisting of 3 activities. The study represents an experiment to evaluate a serious game, with an experimental design that includes a pre-test, the game session, and a post-test; where the evaluation is performed by comparing the pre-post responses and studying whether the way users play the game influences the effects of the game. Simva also allows setting up A/B testing or control group experiments, as different users of the same study can be assigned to different game sessions and/or questionnaires. The module includes a role system and an API to manage the information of the studies. It also allows authorized users to manage and query activities through a scheduler; for example, in studies with multiple activities, they can be configured to be unavailable until the previous activity has been completed.

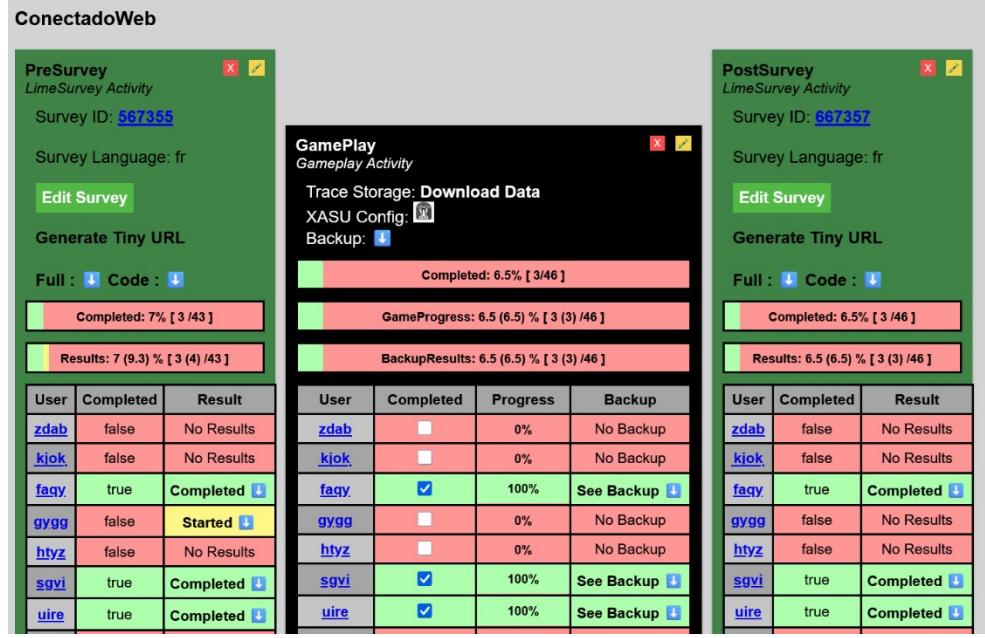


Figure 3: Example of a SIMVA study from the users and activities module of the GLA support system.

In SIMVA there are 2 types of users: basic users, who participate in educational activities and can consult their own progress and information generated during the activity; and advanced users, who can create studies and activities, and grant basic users access to these studies. Advanced users can access all the information generated by the users who participate in one of their studies.

User management is also done in SIMVA, with the possibility to create anonymized user groups and assign them to studies/activities. SIMVA's user management complies with all privacy regulations, including the European GDPR: SIMVA can generate lists of fully anonymous unique identifiers (alphanumeric tokens) that are meant to be handed to participants in a study/activity outside of the system. This way, no personal data is collected, and there is no possibility to link collected data to any participant. SIMVA uses the same anonymous identifiers throughout all activities in a study, linking together all activities from the same user, from gameplay to questionnaire responses, so that further analysis can be performed on player data while maintaining privacy requirements.

Questionnaires are managed through the connection with a survey system, LimeSurvey [17], that allows easy creation and configuration of questionnaires to be included as activities in the studies. Gameplay activities can also be tracked and added into a study. The information on those activities will be gathered thanks to the trackers in the telemetry module.

2.2. Telemetry module

The telemetry module is composed of trackers to capture players' interaction data. A tracker is a software component that deals with the communication of the player interactions (data traces) to a permanent external storage to allow the analysis of behavior and learning outcomes. Currently, the telemetry module is composed of 2 trackers: one implemented for JavaScript (JSTracker) and one for Unity (Xasu). Both trackers support and implement the xAPI standard. As a result, they can be integrated into any JavaScript or Unity game to facilitate the sending out of traces using the xAPI standard. Traces stored in Simva and then processed using the queues capabilities of Kafka.

The JsTracker is designed for JavaScript-based web games. When integrated into a game, it sends player interaction data to the server. It can also create a backup of collected data. To streamline the experimental research process, the game's URL can be added to SIMVA, allowing the scheduler to launch the game at the appropriate stage of the experiment (once the user has completed the pregame

survey) and allowing the scheduler to detect automatically when the game has finished to launch the next stage of the experiment (for example, the postgame survey) when certain conditions are met: the game sends a trace with the verb as “completed”, the object type as “serious-game” and result completion is “true”.

Xasu is a Unity plugin that includes a tracker capable of storing player interaction data either locally or on a server. It offers straightforward tracking functionality and flexible tools for configuring the collection of xAPI traces in a robust manner. Xasu can also send backups of interaction data to the server. To connect Xasu to a specific research study, users can generate and download a Xasu Configuration file from SIMVA. This configuration file provides the necessary settings to Xasu, which automatically launches the game with the correct setup and sends interaction traces to the SIMVA server. In the case of Unity games, the system also provides a complementary plugin, the SIMVA Unity Plugin, to simplify the integration process and allow coordination between the activities of a study. The SIMVA Unity Plugin ensures proper user-session tracking, and helps unify all study components within a single, coherent workflow. It works together with Xasu to automatically handle the connection between the game and the study backend.

2.3. Persistence module

In the system, data persistence is managed in a structured and secure way. All user-generated information (traces of interaction with the game) is stored through a Kafka message queue, which decouples event generation from storage, facilitates system scaling and improves fault tolerance.

The platform uses MinIO, an open-source high performance object storage that emulates Amazon’s S3, as data storage for player interactions. Collected information is saved in MinIO as JSON files, and it is accessible, and a backup of the data is downloadable by the stakeholders in the SIMVA study page. All information is saved using the xAPI format, which then facilitates its connection and reuse by other modules or partners. The information stored is accessible to the creator of the activity, and for players. The information may also be gathered in any available Learning Record Store (LRS), increasing interoperability. An LRS is a system that stores, organizes and allows access to xAPI learning data generated by users while interacting with platforms, games, simulators, or any digital learning environment, in a secure, organized and reusable way. Access to the data is controlled by a system of permissions and authentication, which clearly defines who can submit information to the system (in our case, games using xAPI-SG trackers) and who can access the data (researchers, teachers in SIMVA or in analysis tools). Each issue entity must be authenticated and have explicit permissions to publish events, ensuring the integrity and reliability of the system. This architecture makes it possible to record and access what each user has done, when and in what context.

2.4. Analysis and visualization module

The analysis and visualization module is primarily composed of T-Mon [18], a data analysis and visualization tool specifically designed to process data in xAPI-SG format. T-Mon can read xAPI-SG statements directly from the SIMVA persistence module (via a connected user account) or from an externally uploaded JSON file. T-Mon then analyzes the xAPI-SG data by using the fields (verbs, activity types and extensions) defined in the xAPI-SG format. Those pre-defined fields allow T-Mon to create the analysis and visualizations by default, without any required additional configuration. With the data analyzed, it then generates a default set of visualizations to display the analysis results like showed as an example in Figure 4. This way, T-Mon provides a quick overview of the collected interaction data.



Figure 4: example of a default figure in T-Mon using xAPI-SG data structure.

T-Mon was originally developed as a set of Jupyter Notebooks. Now, it is being migrated to a web-based Python application using Dash, to improve performance and streamline report generation, particularly for large datasets, enabling also the display of the xAPI-SG as a data table. In addition to the default analyses and visualizations included in the visualization module, users have the option to download the data in SIMVA and perform custom analyses or create their visualizations outside of SIMVA using their preferred tools such as Python, R, Excel or Power BI. This flexibility allows stakeholders to adapt the analysis to their reporting needs (pedagogical goals, research questions, etc.), going beyond the standard visualizations offered by the platform.

3. Conclusions and future work

The SIMVA system aims to bridge the gap between serious game design and development and actual game validation in real educational settings. The four modules that compose it (telemetry, users and activities, persistence and visualization) provide full support for game learning analytics from their inception in game interactions (telemetry), to their composition in studies (users and activities), collection (persistence), and their analysis and visualization.

The current version of SIMVA includes a fully developed open-source standard-based version of all four modules. However, it is possible to personalize the different modules or even create new ones from scratch: being open-source and available on GitHub allows an easy and free option to modify or extend them as needed.

The GLA support system aims to simplify all parts of serious games validation and application while complying with the different data privacy regulations, including the EU GDPR. Our goal is to promote the use of GLA when using serious games to expand their use to even more learning scenarios.

Acknowledgements

This work was partially funded by the Ministry of Education (PID2020-119620RB-I00; PID2023-149341OB-I00), and by the Telefonica-Complutense Honorary Chair on Digital Education and Serious Games.

Declaration on Generative AI:

The author(s) have not employed any Generative AI tools.

References

- [1] T. Elias, "Learning Analytics: Definitions, Processes and Potential," 2011, [Online]. Available: <http://learninganalytics.net/LearningAnalyticsDefinitionsProcessesPotential.pdf>
- [2] M. Freire, Á. Serrano-Laguna, B. M. Iglesias, I. Martínez-Ortiz, P. Moreno-Ger, and B. Fernández-Manjón, "Game Learning Analytics: Learning Analytics for Serious Games," in Learning, Design, and Technology, Cham: Springer International Publishing, 2016, pp. 1–29. doi: 10.1007/978-3-319-17727-4_21-1.
- [3] Y. J. Kim, J. A. Ruipérez-Valiente, P. Tan, L. Rosenheck, and E. Klopfer, "Towards a Process to Integrate Learning Analytics and Evidence-Centered Design for Game-based Assessment," Companion Proceedings of the 9th International Conference on Learning Analytics and Knowledge, no. March, pp. 8–10, 2019.
- [4] C.-I. Hwang and S. H. Kim, "How are Serious Games used in the Classroom setting? - Based on the Learning Theory," Indian J Sci Technol, vol. 9, no. 26, Jul. 2016, doi: 10.17485/ijst/2016/v9i26/97307.
- [5] C. Alonso-Fernández, A. Calvo-Morata, M. Freire, I. Martínez-Ortiz, and B. Fernández-Manjón, "Applications of data science to game learning analytics data: A systematic literature review," Comput Educ, vol. 141, p. 103612, Nov. 2019, doi: 10.1016/j.compedu.2019.103612.
- [6] S. A. Alserri, N. A. M. Zin, and T. S. M. T. Wook, "Instrument Validation for Evaluating Serious Game Engagement Model," in 2019 International Conference on Electrical Engineering and Informatics (ICEEI), IEEE, Jul. 2019, pp. 170–175. doi: 10.1109/ICEEI47359.2019.8988873.
- [7] V. Shute, F. Ke, and L. Wang, "Assessment and Adaptation in Games," in Instructional Techniques to Facilitate Learning and Motivation of Serious Games, Cham: Springer International Publishing, 2017, pp. 59–78. doi: 10.1007/978-3-319-39298-1_4.
- [8] Y. Zhonggen, "A Meta-Analysis of Use of Serious Games in Education over a Decade," International Journal of Computer Games Technology, vol. 2019, pp. 1–8, Feb. 2019, doi: 10.1155/2019/4797032.
- [9] F. Bellotti, R. Berta, and A. De Gloria, "Designing Effective Serious Games: Opportunities and Challenges for Research," International Journal of Emerging Technologies in Learning (iJET), vol. 5, no. SI3, p. 22, Nov. 2010, doi: 10.3991/ijet.v5s3.1500.
- [10] S. Judel and U. Schroeder, "EXCALIBUR LA - An Extendable and Scalable Infrastructure Build for Learning Analytics," in 2022 International Conference on Advanced Learning Technologies (ICALT), IEEE, Jul. 2022, pp. 155–157. doi: 10.1109/ICALT55010.2022.00053.
- [11] V. Aleven, J. Sewall, B. M. McLaren, and K. R. Koedinger, "Rapid Authoring of Intelligent Tutors for Real-World and Experimental Use," in Sixth IEEE International Conference on Advanced Learning Technologies (ICALT'06), IEEE, pp. 847–851. doi: 10.1109/ICALT.2006.1652575.
- [12] N. Joudieh, W. van der Aalst, R. Champagnat, M. Rabah, and S. Nowakowski, "Moodle2EventLog: A Tool for Pedagogically-Driven Log Enrichment and Analysis," in Proceedings of the 17th International Conference on Computer Supported Education, SCITEPRESS - Science and Technology Publications, 2025, pp. 452–463. doi: 10.5220/0013327300003932.
- [13] B. Heinemann, M. Ehlenz, S. Görzen, and U. Schroeder, "xAPI Made Easy: A Learning Analytics Infrastructure for Interdisciplinary Projects," International Journal of Online and Biomedical Engineering (iJOE), vol. 18, no. 14, pp. 99–113, Nov. 2022, doi: 10.3991/ijoe.v18i14.35079.
- [14] L. Jean Justice and A. D. Ritzhaupt, "Identifying the Barriers to Games and Simulations in Education," Journal of Educational Technology Systems, vol. 44, no. 1, pp. 86–125, Sep. 2015, doi: 10.1177/0047239515588161.
- [15] C. Alonso-Fernández, A. Calvo-Morata, M. Freire, I. Martínez-Ortiz, and B. Fernández-Manjón, "Simplifying the Validation and Application of Games with Simva," in Emerging Technologies for Education, 2020, pp. 337–346. doi: 10.1007/978-3-030-38778-5_37.
- [16] Á. Serrano-Laguna, I. Martínez-Ortiz, J. Haag, D. Regan, A. Johnson, and B. Fernández-Manjón, "Applying standards to systematize learning analytics in serious games," Comput Stand Interfaces, vol. 50, pp. 116–123, Feb. 2017, doi: 10.1016/j.csi.2016.09.014.
- [17] The LimeSurvey Project Team, "LimeSurvey." [Online]. Available: <https://www.limesurvey.org/>
- [18] C. Alonso-Fernández, A. Calvo-Morata, M. Freire, I. Martínez-Ortiz, and B. Fernández-Manjón, "Game Learning Analytics: Blending visual and data mining techniques to improve serious games and to better understand players learning," Journal of Learning Analytics, vol. 9, no. 3, pp. 32–49, Dec. 2022, doi: 10.18608/jla.2022.7633.